Probability Practice

Jason Antal

2024-08-18

QUESTION: Part A. Visitors to your website are asked to answer a single survey question before they get access to the content on the page. Among all of the users, there are two categories: Random Clicker (RC), and Truthful Clicker (TC). There are two possible answers to the survey: yes and no. Random clickers would click either one with equal probability. You are also giving the information that the expected fraction of random clickers is 0.3. After a trial period, you get the following survey results: 65% said Yes and 35% said No. What fraction of people who are truthful clickers answered yes? Hint: use the rule of total probability.

```
knitr::opts_chunk$set(echo = TRUE)

p_rc <- 0.3  # Probability of being a Random Clicker
p_yes <- 0.65  # Overall probability of answering Yes
p_yes_rc <- 0.5  # Random Clickers have 50% chance of answering Yes

p_yes_tc <- (p_yes - p_yes_rc * p_rc) / (1 - p_rc)

#ANSWER:
cat("The percent of Truthful Clickers who answered Yes is:", round(p_yes_tc,4)*100, "%.", '\n',"As a fra
```

```
## The percent of Truthful Clickers who answered Yes is: 71.43 %.
##  As a fraction, this means roughly 7/10 of truthful clickers answered yes.
```

QUESTION: Part B. Imagine a medical test for a disease with the following two attributes:

```
The sensitivity is about 0.993. That is, if someone has the disease, there is a probability of 0.993 tha
The specificity is about 0.9999. This means that if someone doesn't have the disease, there is probabili
In the general population, incidence of the disease is reasonably rare: about 0.0025% of all people have
```

Suppose someone tests positive. What is the probability that they have the disease?

```
sensitivity <- 0.993  # P(Positive | Disease)
specificity <- 0.9999  # P(Negative | No Disease)
prevalence <- 0.000025  # P(Disease)

# Calculate P(Positive)
p_positive <- sensitivity * prevalence + (1 - specificity) * (1 - prevalence)

# Calculate P(Disease | Positive)
p_disease_given_positive <- (sensitivity * prevalence) / p_positive

# ANSWER
cat('The probability that someone who tests positive for the disease actually has the disease is', round
```

```
## The probability that someone who tests positive for the disease actually has the disease is 19.89 %.
```